

Remarks

Claims 1-7 and 9-10 are pending in this application. Claims 1-2 and 10 stand rejected under 35 U.S.C. 103(a) as being unpatentable over Huang et al. (U.S. Patent No. 5,913,193) in view of Hata et al. (U.S. Patent No. 5,878,393). Claims 3-5 and 9 stand rejected under 35 U.S.C. 103(a) as being unpatentable over Huang et al. in view of Hata et al. and further in view of Trader et al. (U.S. Patent No. 5,832,432). Claims 6 and 7 stand rejected under 35 U.S.C. 103(a) as being unpatentable over Huang et al. in view of Hata et al. and further in view of Holm et al. (U.S. Patent No. 5,850,629).

Claim 1 has been amended to further recite that the recordings for a single inflection of a single speech item are a limited set of recordings that represent a limited set of ligatures with adjacent speech items including only recordings having a vowel at either end and recordings having no surrounding ligature distortions.

This aspect of the invention is exemplified in the specification at page 30, line 27-page 32, line 35, among other places. More specifically, page 32, lines 24-35 specifically convey the concept of utilizing a limited set of recordings that represent a limited set of ligatures with adjacent speech items including only certain recordings.

With regard to claim 1, Huang fails to describe multiple voice recordings in the digital voice library that correspond to a single inflection of a single speech item, for a plurality of inflections of a plurality of speech items, that represent various ligatures for the single inflection of the single speech item with adjacent speech items, and the additional added claim limitations.

Huang does describe a method and system of run time acoustic unit selection for speech synthesis. Huang discusses acoustic unit selection for speech synthesis and attempts to minimize the spectral distortion between boundaries of adjacent instances. Huang discusses a training phase where the highest probability instances representing diverse phonetic contexts

are chosen. Col. 2, lines 65-67. Huang further describes that the naturalness of synthesized speech is improved by providing multiple instances of an acoustic unit, such as a diphone. Col. 9, lines 58-60.

Huang only describes the creation of a robust library with multiple instances of an acoustical unit but fails to suggest the specific techniques set forth by independent claim 1 involving inflections and ligatures.

In making the rejection, the Examiner relies on Hata as a secondary reference.

Hata describes a concatenative reading system. Hata does describe the broad concept of pronouncing words differently, based on the ending phoneme of the preceding word and the beginning phoneme of the following words. The Examiner has directed Applicant's attention to Hata, Col. 4, lines 28-55. This portion of Hata does describe dictionary samples including a different sample for each pitch contour (lines 28-31). Also described is the concept of pronouncing phonemes differently, depending on what sounds precede and follow the phoneme.

Hata fails to recognize or suggest using a limited set of recordings that represent a limited set of ligatures with adjacent speech items including only recordings having a vowel at either end and recordings having no surrounding distortions as recited by claim 1 but instead Hata only generally discusses concatenation and pronunciation. Applicant recognizes the advantages of utilizing the limited set of recordings, and Hata makes no suggestion of this aspect of the claimed method. Accordingly, Hata fails to overcome the deficiency of Huang.

With regard to Trader, relied upon by the Examiner as a further secondary reference, Trader describes a method for converting a text classified ad to a natural sounding audio ad. Ads from a print media ad database are downloaded to a host processor that identifies relevant ads and converts the abbreviated text of the print ad to an expanded version of the ad. The words and phrases of the expanded version are parsed and stored in appropriate

fields of a relational database. A sequence play list of audio file numbers corresponding to the words and phrases contained in a database record is created along with glue words and phrases that are added to produce a more natural sounding audio. Trader fails to describe or suggest, as specifically recited in claim 1, establishing multiple voice recordings in a digital voice library that correspond to a single inflection of a single speech item, for a plurality of inflections of a plurality of speech items, that represent various ligatures for the single inflection of the single speech item with adjacent speech items, and the additional added claim limitations.

As noted above, Huang has significant shortcomings. As well, Hata and Trader have these shortcomings and do not overcome the deficiency of Huang. With regard to Holm, Holm does not overcome the shortcomings noted above with respect to Huang, Hata, and Trader. Accordingly, claims 1-7 and 9-10 are believed to be in condition for allowance and such action is respectfully requested.

Please charge any fees or credit any overpayments as a result of the filing of this paper to our Deposit Account No. 02-3978.

Respectfully submitted,

ELIOT M. CASE ET AL.

By: /Jeremy J. Curcuri/
Jeremy J. Curcuri
Reg. No. 42,454
Attorney for Applicants

Date: December 5, 2007

BROOKS KUSHMAN P.C.
1000 Town Center, 22nd Floor
Southfield, MI 48075-1238
Phone: 248-358-4400
Fax: 248-358-3351